DOCUMENT RESUME

ED 211 981 CS 206 700

AUTHOR	Witte, Stephen P.; And Others
TITLE	The Empirical Development of an Instrument for Reporting Course and Teacher Effectiveness in College Writing Classes. Technical Report Number 3.
INSTITUTION	Texas Univ., Austin.
SPONS AGENCY	Fund for the Improvement of Postsecondary Education (ED), Washington, D.C.
PUB DATE	31 Aug 81
GRANT	G008005896
NOTE	54p.; Not available in paper copy due to marginal legibility of original document. Prepared through the Writing Program Assessment Project.
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	*Course Evaluation; Evaluation Methods; Higher Education; *Measures (Individuals); Questionnaires; Rating Scales; *Student Evaluation of Teacher Performance; *Teacher Effectiveness; *Test Construction; *Writing Evaluation; Writing Instruction
IDENTIFIERS	Writing Program Assessment Project

ABSTRACT

As part of a three-year project that approached writing program evaluation through writing itself, the teaching of writing, and the goals of the writing programs, this report describes the development of an instrument for reporting course and teacher effectiveness in college writing classes. The report describes the instrument, which was developed to help provide teachers with adequate information about the students' perceptions of their writing courses and their teaching and to provide an easily scored, reliable, and valid evaluation form for eliciting students' judgments about specific aspects of writing courses and the teaching of writing. Sections in the report discuss (1) the effectiveness of student rating scales; (2) the stages of development for the evaluation instrument--item collection and generation, testing of the preliminary instrument, testing and revising the second version of the instrument, and testing the revised version; (3) analysis of the two forms of the teacher and course effectiveness questionnaire, including a description of the sample population and results of the factor analysis; (4) canonical correlations between the two parts of the teacher and course effectiveness questionnaire; and (5) the scoring, uses, and limitations of the instrument. Appendixes include both the long and the short forms of the instrument for evaluating course and teacher effectiveness. (HOD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED211981

WRITING PROGRAM ASSESSMENT PROJECT
Technical Report Number 3

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

THE EMPIRICAL DEVELOPMENT OF AN INSTRUMENT FOR REPORTING COURSE AND TEACHER EFFECTIVENESS IN COLLEGE WRITING CLASSES

Stephen P. Witte
John A. Daly
Lester Faigley
William R. Koch

BEST AVAILABLE COPY

Jointly Funded by The Fund for The Improvement of Post-Secondary Education and The University of Texas

CS206700

TECHNICAL REPORT NO. 3

FUND FOR THE IMPROVEMENT OF POSTSECONDARY EDUCATION
GRANT NO. G008005896

THE EMPIRICAL DEVELOPMENT OF AN INSTRUMENT FOR REPORTING COURSE AND TEACHER EFFECTIVENESS IN COLLEGE WRITING CLASSES

Stephen P. Witte,
John A. Daly, Lester Faigley
and William R. Koch

The Writing Program Assessment Project
GRG 106-A
The University of Texas at Austin
Austin, Texas 78712

August 31, 1981

The research reported herein was begun under a grant from The University Research Institute at The University of Texas at Austin and continued under a grant from the Fund for the Improvement of Postsecondary Education with supplemental support from The University of Texas at Austin. The opinions expressed in this document do not necessarily reflect the positions or policies of the funding agencies.

TABLE OF CONTENTS

TABLE OF CONTENTS i

CHAPTER I: BACKGROUND OF THE PRESENT STUDY 1

I.1. INTRODUCTION 1

I.2. EFFECTIVENESS INSTRUMENTS 1

CHAPTER II: DEVELOPMENT OF THE EVALUATION INSTRUMENT 7

II.1. INTRODUCTION 7

II.2. STAGES OF DEVELOPMENT 7

II.2.1. Stage 1--Item Collection and Generation . 8
II.2.2. Stage 2--Testing of the Preliminary Instrument 8
II.2.3. Stage 3--Testing and Revising the Second Version of the Instrument 9
II.2.4. Stage 4--Testing the Revised Version . . . 11

CHAPTER III: ANALYSES OF THE TWO FORMS OF THE TEACHER AND COURSE EFFECTIVENESS QUESTIONNAIRE 13

III.1. INTRODUCTION 13

III.2. DESCRIPTION OF THE SAMPLE POPULATION 13

III.3. RESULTS OF THE FACTOR ANALYSIS 15

III.3.1. A Two-Factor Solution for Part I of the Questionnaire 17
III.3.2. Coefficient Alpha Reliabilities for Part I 19
III.3.3. A Nine-Factor Solution for Part II of the Questionnaire 19
III.3.4. Coefficient Alpha Reliabilities for Part II 28

III.4. CANONICAL CORRELATIONS BETWEEN THE TWO PARTS OF THE TEACHER AND COURSE EFFECTIVENESS QUESTIONNAIRE 28

III.4.1. Introduction 29
III.4.2. Results of the Canonical Correlation

Analysis 29

III.5. THE SHORT FORM OF THE COURSE AND TEACHER EFFECTIVENESS INSTRUMENT 32

III.5.1. Development of the Short Form 32
III.5.2. A Two-Factor Solution for the Short Form 33
III.5.3. Coefficient Alpha Reliabilities for the Short Form of the Instrument . . . 35

CHAPTER IV: THE SCORING, USES, AND LIMITATIONS OF THE INSTRUMENT 37

IV.1. INTRODUCTION 37

IV.2. SCORING THE LONG AND SHORT FORMS 37

IV.3. RECOMMENDED USES OF THE INSTRUMENT 38

IV.4. CAUTIONS REGARDING THE USE OF THE INSTRUMENT 39

REFERENCES 43

APPENDIX A: INSTRUMENT FOR EVALUATING COURSE AND TEACHER EFFECTIVENESS: LONG FORM 47

APPENDIX B: INSTRUMENT FOR EVALUATING COURSE AND TEACHER EFFECTIVENESS: SHORT FORM 53

CHAPTER I

BACKGROUND OF THE PRESENT STUDY

I. 1. INTRODUCTION

The present study was undertaken for two reasons. First, we believe that improving the teaching of writing in colleges and universities depends, in part, on providing teachers with accurate information about the students' perceptions of their writing courses and their teaching. Second, we undertook the present study because we were unable to identify an easily scored, reliable, and valid evaluation form which elicited student judgments about specific aspects of writing courses and the teaching of writing. In fact, the only attempt that we know of to produce such an instrument is the ongoing one of the Committee on Teaching and Its Evaluation in Composition associated with the Conference on College Composition and Communication (Larson et al., 1981). While the efforts of that Committee are laudable, the instrument they propose has yet to be tested.

The instrument described in this report has been tested in a number of different colleges and universities. It is an instrument which we believe meets the criteria of reliability, validity, and ease of scoring. In addition, the instrument was designed to give college teachers of writing evaluative information about very specific aspects of teaching and composition courses.

I. 2. EFFECTIVENESS INSTRUMENTS

Since the early work on the Purdue Rating Scale (Shock, Kelly, & Remmers, 1927), considerable attention has been given to the development of standardized or semi-standardized student rating scales of faculty, and a multitude of such instruments are currently in use (see Costin, Greenough, & Menges, 1971). The literature on faculty evaluation in general and the use of student rating scales in particular is large and on several crucial issues, contradictory.

A number of attempts have been made to establish or to argue the validity of such scales (e.g., Eble, 1970; Costin, Greenough, & Menges, 1971; Frey, 1974; Tuchman, 1976; and Whitely & Doyle, 1979), but the issue seems far from settled (Marsh & Overall, 1980; Powell, 1978). Indeed, the lack of agreement over the

nature of effective teaching are ever ways to assess it has led to the production of the large number of different instruments. As Cromack and Eggleston (1979) point out, "the failure to establish a universally accepted criterion of teaching effectiveness has resulted in a proliferation of rating forms and strategies" (p. 8). Such a proliferation would probably not have resulted if educators agreed on what good teaching is and on how to assess it. An acceptable definition of good teaching or what good teachers do must precede any valid measure of teaching.

Despite the widespread use of rating instruments or scales, they may be only marginally reliable in distinguishing among teachers of varying effectiveness. Newton (1979), for example, notes that "Evaluation systems can distinguish very bad (usually few under existing evaluation procedures) from the vast majority of teachers; but few systems separate average from superior or superior from excellent" (p. 45).

Questions have also been raised about the value of the feedback teachers receive from such instruments. James Kulik (1976), addressing specifically the weaknesses of standardized rating scales, suggests that their value is limited. Because they do not give any clear indication of why students responded as they did, they give teachers little direction on what to do to improve their teaching, or ratings. Nevertheless, a recent survey of 134 institutions indicates that student ratings of faculty is the most commonly used form of faculty evaluation (Centra, 1979); and a second survey--this one of college and university writing program directors--reports that student ratings of faculty are the most frequently used form of evaluating faculty in college composition courses (Witte, Meyer, Miller, & Faigley, 1981).

On the issue of whether the feedback teachers receive from such instruments actually improves teaching, the literature is ambiguous. Centra's (1973, 1977) research, for example, led to the conclusion that student ratings of faculty do effect improvements in teaching, a conclusion earlier voiced by Tuchman and Oliver (1968). Kulik and McKeachie (1975), however, argue that evidence of teaching improvement as a result of information garnered from student ratings is far from convincing.

The unresolved questions and the unsettled issues notwithstanding, standardized or semi-standardized instruments designed to elicit student evaluations of college courses and college teaching are being used widely in postsecondary institutions. Viewed in the most favorable light, this use suggests that these institutions are more interested than ever before in teacher effectiveness, that they are more committed to hiring and retaining faculty who are effective classroom teachers, not just capable researchers and knowledgeable experts in their fields. At many colleges and universities, the administration of such instruments is mandatory for all faculty. At other institutions, it is mandatory for temporary and probationary or tenure-seeking faculty. And at most

institutions, the administration of such instruments is recommended for all faculty. Such instruments provide the bases of administrative decisions regarding salary, tenure, and promotion (Thorne, 1980; O'Hanlon & Mortenson, 1980), as well as providing feedback to teachers about their courses and their teaching.

The use of student-completed forms for rating the teaching performances of faculty and the courses they teach is usually accompanied by several problems. Some of these problems are conceptual ones; these conceptual problems often lead to methodological ones.

From a conceptual standpoint, the design of the instruments and their intended uses are of paramount importance. Such instruments may be poorly designed and used for purposes for which they were neither intended nor validated. Most instruments which elicit anonymous evaluations of faculty and the courses they teach have been constructed by committees of faculty and students who have considerable intuitive knowledge of effective teaching and effective courses but who have little or no knowledge of how to validate their intuitons by testing the instruments they design. Generating what appear to be good items for a teacher/course evaluation instrument is not sufficient: items have to be tested repeatedly to determine whether they are capable of distinguishing between effective and ineffective courses and teachers. Typically, such instruments are designed to provide feedback to teachers about their teaching and about their courses. This distinction between the course itself and the teaching of it is an important distinction, because perceptions of one may very well affect perceptions of the other and because most such instruments contain both types of items. It would be appropriate, therefore, to separate items on course content from items on instruction and to obtain separate scores for each set. Yet total scores are often determined by summing responses across items regardless of focus. Thus what is a conceptual problem at one level is a methodological problem at another.

Another conceptual problem reflected in current methodology is that the same teacher evaluation instruments are used in courses which differ fundamentally in kind, even though the courses may be offered within the same college or even within the same department. It is not uncommon to find that the same form used to evaluate teaching in literature courses is also used to evaluate teaching in composition courses, in a course on sociological research methods, in a Russian history course, in an art history course, in a course on symbolic logic, and in a chemistry laboratory course. Such a procedure assumes that teaching, say, research methods in sociology is no different than teaching American literature, that teaching graduate students in a Russian history course is no different than teaching freshmen in a basic writing course.

Certain aspects of teaching most assuredly cross

disciplinary boundaries and levels of instruction: students in all disciplines at all levels like to be treated as human beings, like to have assignments returned to them within a reasonable time, like to be able to talk about courses with instructors. However, in many particulars, teachers teach courses at different curricular levels in different ways, teach students of different ability levels differently, teach the content of one discipline differently than they teach the content of another. We believe that it is necessary to accommodate these contextual differences in any attempt to evaluate teaching.

While it is possible to construct instruments capable of assessing teaching and courses across disciplines and curricular levels, the very fact that such instruments must ignore the peculiar logics or idiosyncracies of different disciplines and the demands of varied student populations decreases the amount of information the instruments can provide a teacher and significantly alters the nature of that information. Instruments used in a variety of courses at a number of curricular levels may provide the bases for interesting comparisons across courses, colleges, or departments; but the information which reaches the teacher's hands is only generally and, frequently, only minimally useful. Such instruments will not tell teachers anything specific about how to improve their teaching during the next semester.

Another problem has to do with the time at which such instruments are administered. Typically, student evaluations of faculty and courses are conducted near the end of a semester or quarter. The result is evaluation data which is summative in nature. Several problems are associated with conducting student-based evaluations of faculty and courses at or near the end of a semester or quarter.

First, the end of a term is that time when students are most anxious about their own performances in classes. Often they are apprehensive about the pending final examination or grading of the exam, about the term paper or project that must be submitted or evaluated at the end of the term. Often they are simply tired from working hard in all their courses or from carrying the pressure that accompanies the endings of terms for most students. It is possible that the anxiety which accompanies the ends of semesters or quarters for most students can affect their judgments of either their teachers or their courses.

Second, conducting student evaluations of faculty and courses at or near the end of a term assumes that (1) that time is the best (and often the only) time to evaluate teachers and courses and (2) the evaluation data collected at or near the end of a term are the most reliable and the most valid data. Both of these assumptions can be challenged. Receiving the results of an evaluation after a course is over can do nothing to improve the teaching of that course until the subsequent term. It may be that a series of formative evaluations conducted at various times during the semester or quarter is preferable to a single

summative evaluation at or near the end of the term. Evaluations not influenced by end-of-semester pressures or anxieties may better indicate the quality of the teaching and the value of the course. Furthermore, the feedback would be timely enough for the teacher to modify instruction while the students can benefit from the changes, and the students would be able to see the effects of their interim evaluations as the course progresses. If, however, it is necessary and desirable to collect post-term evaluations of teachers and courses, perhaps evaluation data would provide more reliable and more valid estimates of teachers and courses if collected sometime after the course were completed.

In the present monograph, we report on the development of an instrument which we believe overcomes the weaknesses of rating scales currently used to evaluate composition instruction in colleges and universities. Those readers interested only in using the instrument in its long or short form will find the forms in APPENDIX A and APPENDIX B and directions for administering them in Chapter IV. Chapter II and Chapter III are included for the benefit of those interested in the development and testing of the instrument and in the factor structures underlying the long and short forms.

CHAPTER II

DEVELOPMENT OF THE EVALUATION INSTRUMENT

II. 1. INTRODUCTION

The instrument for evaluating the effectiveness of teachers in college composition courses was developed in four stages.

The first stage involved collecting or generating a large pool of statements to which students could respond about their composition teachers and classes. During the second phase, a large number of students completed pilot questionnaires that included the statements formulated during the first stage of the investigation. The purpose of this second stage was to generate an inter-item correlation matrix that would allow us to examine the relationships among items and subsequently to reduce the number and refine the wording of the items retained. The third stage involved administering the revised version of the instrument to a large number of students in different colleges and universities, and then factor analyzing the inter-item correlation matrix generated from the student responses.

During the fourth stage--which is the primary focus of this report--the instrument used during the third stage was modified on the basis of the factor analyses performed during that earlier stage. Subsequently, the revised instrument was administered to students in nine colleges and universities. From these student responses, another correlation matrix was generated and factor analyzed. From the analyses performed during the fourth stage, an empirically derived instrument for reporting student reactions to composition teaching and composition courses was developed. The long form of the instrument consists of two parts. The first and shorter part consists of items designed to assess the more general aspects of composition teaching and composition courses. The larger second part consists of items designed to assess specific aspects of writing instruction and courses. Factor analyses of the two parts were conducted and canonical correlation analysis was used to determine the relationship between the two parts of the instrument. Subsequently, a short form of the instrument was developed and submitted to factor analyses. Samples of the long and short forms of the instrument appear in APPENDIX A and APPENDIX B, respectively.

II. 2. STAGES OF DEVELOPMENT

Each of the four stages is described briefly in the sections

which follow. The analytical aspects of Stage 4 are detailed in Chapter III.

II. 2. 1. Stage 1--Item Collection and Generation

The preliminary instrument was developed in three steps. First, a number of undergraduate students enrolled in writing courses at both the lower-division and upper-division levels, graduate teaching assistants, composition teachers from seven institutions, and four writing programs directors, were asked to describe the major characteristics that distinguish good and poor composition teaching and courses at the college level. These statements were used to identify a variety of hypothetical dimensions of composition teaching for which items could be collected or generated. Second, items reflecting each of these dimensions were generated. These items consisted of statements made by college composition students, teachers, and writing program directors or of statements--based on information about writing courses given by the three groups--written by the investigators. Because these items were developed from statements made by composition students, teachers, and writing program directors about college composition courses and teaching, we believe the pool of statements represented content-valid dimensions of college composition classes. Third, the investigators divided these statements into two groups, one containing items describing composition instruction and courses and the other containing items evaluating composition instruction and courses.

These two major categories reflected the basic differences in the statements and information supplied by the composition students, teachers, and writing program directors we consulted in developing the items. These two broad categories were thought to be potentially useful for future work. In all, 207 statements about composition instruction were included on the preliminary instrument. The categories of descriptive items and evaluative items provided the basis for what was a two-part preliminary instrument. With the descriptive items, we hoped to identify the frequency with which teachers exhibited certain teaching behaviors in the classroom; with the evaluative items, we hoped to determine how effective certain behaviors were. Descriptive items were distributed randomly within one part, and evaluative items were distributed randomly within the second part. Each hypothetical dimension was represented on one or the other of the two parts of the preliminary instrument.

II. 2. 2. Stage 2--Testing of the Preliminary Instrument

The preliminary instrument was administered to over 400

undergraduate students enrolled in composition courses at a large public university, a private four-year institution, and a public community college. These students were asked to use the statements on the instrument to assess their composition teachers and classes. In responding to the descriptive items, students used a five-point Likert-type scale with ranges from "very often" to "never" to indicate the frequency with which a particular teacher behavior occurred in their classes. In responding to the evaluative statements, the students again used a five-point Likert-type scale, this one reflecting verbal extremes from "strongly agree" to "strongly disagree."

The students' responses to the 207-item questionnaire were coded and descriptive statistics were computed for them. A close examination of the responses provided the bases for subsequent revision of the instrument. The major issues in the analyses were the variability in student responses to each item and the interrelationships among the items. If every student responded to an item in approximately the same way, the item was either rewritten or deleted from the questionnaire because the item was not functioning to distinguish teacher or course effectiveness. We wanted items that would elicit variation in responses for different teachers and different classes. Groups of items that had high correlations were carefully examined and needlessly redundant items were removed. Items thought to be tapping important dimensions but which had negligible associations were carefully considered. Sometimes such items were rewritten; sometimes additional items were composed so that the dimension might be better assessed. In contrast, if an item seemed to be a poor one in terms of content or clarity, it was deleted from the pool of statements. In addition to relying on statistical analyses of the student responses in determining how the questionnaire ought to be revised, we examined statements which students and teachers had made about the instrument in general or about particular items included on the preliminary version. These statements, coupled with our statistical analyses of the responses, indicated the perceived clarity, importance, and validity of the items.

At the end of the second stage, we had a significantly shorter instrument--containing only 132 items as compared to the 207 items of the first version--that seemed clearly written and useful for the assessment of composition teaching. Of the 132 total items, 48 were descriptive statements included on one part of the questionnaire and 84 were evaluative statements and included on the second part. Given the sources of the statements, their content validity was assumed. These items, distributed within a two-part revised questionnaire, were the focus of the third stage of the investigation.

II. 2. 3. Stage 3--Testing and Revising the Second Version of the Instrument

During the third stage, we tested the second version of the instrument. Over 500 students from four institutions participated--some from a community college, some from a private four-year institution, and some from two large public universities. All undergraduate curricular levels were represented in the sample, although most of the students were lower-division students about half of whom were enrolled in their first course in college writing and about half in their second.

As we reported elsewhere (Witte, Daly, & Faigley, 1980), the responses to each of the two parts of the questionnaire were factor analyzed separately. The analysis performed was a principal components factor analysis with orthogonal rotation as determined by Kaiser's varimax criterion (Kaiser, 1958; Morrison, 1967). An eigenvalue of 1.0 was set for termination of factor extraction, a criterion commonly used to determine the number of factors present when principal components analysis is performed (Kaiser, 1974; Kim & Mueller, 1978). An item was deemed statistically satisfactory if it met one of two criteria: (1) if its factor loading was greater than 0.60 and if it loaded no higher on any other factor than 0.40 or (2) if its factor loading was greater than 0.50 and if the item loaded no higher than 0.30 on any other factor.

It has been said that factor analysis is as much an art as it is a science. In this case, neither the art nor the science of factor analysis served us as well as we had hoped it would. Although we examined many possible solutions--from two-factor solutions to 28-factor solutions--for the two parts of the questionnaire, in neither instance were we able to identify a completely satisfactory solution. However, we opted for an 11-factor solution to the descriptive part of the questionnaire and a 21-factor solution for the evaluative part. The 11-factor solution for the descriptive items retained 30 of the original 48 items and accounted for 54% of the variance within the responses. A dominant factor, which we called "Types of Writing," alone accounted for 19.3% of the variance. Seven of the 30 items loaded satisfactorily on that factor. The 21-factor solution for the evaluative part of the questionnaire accounted for 62.5% of the variance and retained 63 of the original 94 evaluative statements. A very general factor--which we called "Conduct of the Class"--accounted for 27.4% of the variance within the responses to the evaluative statements. Of the 63 evaluative statements retained in the 21-factor solution, 29 loaded on this factor.

As we have noted, the factor solutions for the second version of the questionnaire were not altogether satisfactory, both because of the large number of factors and the difficulty experienced in naming a few of the factors in each of the two adopted solutions. In addition, a number of students who completed the questionnaire and a number of teachers who administered it questioned the wisdom of having both descriptive and evaluative statements on the same instrument. Several of the teachers made comments such as the following, which one teacher

returned with her students' completed questionnaires:

> . . . So what if I don't teach description "very often"? My students already know how to write descriptions when they enter my class. Of course, I'm not going to teach them "very often" or even "sometimes" what they already know. I just don't see how your descriptive items can give me any very useful information about my teaching. I know what I do, and I usually know how often I do it. What I want to know is how well I teach description on those rare or those frequent occasions when I do so.

Statements such as these from teachers whose opinions we respect and whose teaching we admire forced us to rethink the nature of the questionnaire as it then it existed. Given such statements and some factors which were at best ambiguous and at worst uninterpretable, we decided to make the entire questionnaire an evaluative one.

To so change the nature of the questionnaire, we recast as evaluative statements the 30 descriptive items retained in the 11-factor solution. Thus a descriptive item which originally read, "The instructor provides students with models of good writing," became, "The instructor <u>is good about providing</u> students with models of good writing." In addition to converting the descriptive statements to evaluative ones, we added six new items. These six items were designed to evaluate general rather than specific aspects of writing instruction.

At the end of the third stage, then, we had a new instrument. This new instrument consisted of 99 statements designed to evaluate both general aspects of teaching and courses <u>and</u> particular aspects of composition teaching and courses.

II. 2. 4. <u>Stage 4--Testing the Revised Version</u>

During the fourth stage of the investigation, 1,552 students evaluated their composition teachers and classes with the revised two-part questionnaire. One part contained 20 general evaluation items, mostly items of the kind which appear on other instruments used to evaluate teaching across disciplines and across different student populations. One such item was "As a student in this course, I am treated fairly." The second part of the questionnaire contained 79 more specific evaluative items. One of these items was, "The instructor is good about teaching me to use new words that help to make my writing more precise." The students responded to these evaluative statements along a five-point Likert-type continuum having a range of possible responses from "strongly agree" to "strongly disagree." In

addition, a sixth possible response category was provided which allowed students to specify those evaluative statements they believed "not applicable" to their particular composition classes or teachers. Students also supplied demographical and biographical information about themselves and their previous educational experiences.

The student responses to the revised version were then submitted to factor analyses. Separate factor analyses were performed on the two parts of the questionnaire. For the first part of the questionnaire--the part containing 20 general evaluative statements--we chose a two-factor solution which retained 18 of the original 20 items. For the second part which dealt with specific aspects of composition teaching and composition courses, we chose a nine-factor solution. This solution retained 62 of the original 79 evaluative statements. Subsequently, we performed a canonical correlation analysis to determine the relationship between Part I and Part II of the questionnaire, which now consisted of 18 and 62 items, respectively. With this analysis we observed six statistically significant canonical correlations between the two parts of the questionnaire. In addition, we developed from the two parts of the long form of the questionnaire a 21-item short form. This form was also factor analyzed. These analyses revealed that a general two-factor solution was the most interpretable solution.

These analyses, together with interpretations of the factor solutions and the canonical correlations, are the subjects of the following chapter.

CHAPTER III

ANALYSES OF THE TWO FORMS OF THE TEACHER AND COURSE EFFECTIVENESS QUESTIONNAIRE

III. 1. INTRODUCTION

The present chapter deals with the more analytic and interpretative aspects of the present study. It treats the following major topics: (1) the sample population on which the analyses are based; (2) the factor solutions of the two parts of the long questionnaire and our interpretation of those solutions; (3) the canonical correlations between the two parts of the questionnaire; and (4) the development of the 21-item short form of the questionnaire.

III. 2. DESCRIPTION OF THE SAMPLE POPULATION

Although some 17 institutions participated in the various stages of developing the questionnaire, the student responses analyzed during the fourth stage came from nine institutions. Using the taxonomies employed by the National Center for Educational Statistic (Dearman, 1980; Grant & Eidon, 1980) and published descriptions of these institutions (Biesel et al., 1979; Watts & Hunter, 1979; Cass & Birnbaum, 1979; Matheson, 1980), we classified the institutions by type and source of funding. Four of the institutions receive their primary funding from private sources, and five receive it from public. Two of the institutions are two-year colleges; three are four-year institutions; and four are universities. Both of the two-year institutions are public institutions; one of the four-year institutions is a public school and the other two are private; one of the universities is a private institution, while the remaining three are public. Student enrollments in the nine institutions range from approximately 2,500 to approximately 49,000. Two of the schools are located in the Northeast, one in the Southeast, five in the Southwest, and one in the West. We believe that the institutions adequately represent the spectrum of undergraduate institutions and student bodies.

A total of 1,552 usable student responses to the questionnaire were collected and analyzed, although the N's for the various test items vary somewhat. Table III.1 shows the distribution of the 1,552 students across types of institutions and the principal sources of funding for those institutions.

TYPE OF INSTITUTION	PRIVATE N	PRIVATE %	PUBLIC N	PUBLIC %	TOTALS N	TOTALS %
Two-Year	---	-----	112	7.22	112	7.22
Four-Year	283	18.23	88	5.67	371	23.90
University	195	12.56	874	56.32	1069	68.88
TOTALS	478	30.79	1074	69.21	1552	100.00

Table III.1. Distribution of Students Across Institutional Types.

As Table III.1 shows, 7.22% of the students were from two-year colleges, 23.67% from four-year institutions, and 69.11% from universities.

These 1,552 students were enrolled in 80 different composition classes. The distribution of these classes across types of institutions is presented in Table III.2.

TYPE OF INSTITUTION	PRIVATE N	PRIVATE %	PUBLIC N	PUBLIC %	TOTALS N	TOTALS %
Two-Year	--	-----	8	10.00	8	10.00
Four-Year	18	22.50	3	3.75	21	26.25
University	9	11.25	42	52.50	51	63.75
TOTALS	27	33.75	53	66.25	80	100.00

Table III.2. Distribution of Composition Classes in Sample Across Institutional Types.

Table III.2 shows that 10% of the students were enrolled in classes taught in two-year colleges, 26.25% in classes taught in four-year institutions, and 65.75% in classes taught in universities.

These 80 classes consisted of both freshman and nonfreshman composition classes. Some of the freshman classes were developmental or remedial classes, although we cannot tell from

the information received exactly how many. Most classes, however, were either first- or second-semester/quarter "regular" freshman writing classes. Of the 80 classes, 25% were nonfreshman classes. Of these 20 nonfreshman classes, eight were sophomore-level courses which were taken after the students completed a freshman course or two freshman courses in writing. The remaining 12 nonfreshman classes were either business writing classes or advanced expository writing classes taught at the junior level. The distribution of classes across freshman and nonfreshman courses is summarized in Table III.3.

TYPE OF INSTITUTION	FRESHMAN N	FRESHMAN %	NONFRESHMAN N	NONFRESHMAN %	TOTAL N	TOTAL %
Two-Year						
Private	--	-----	--	-----	--	-----
Public	8	10.00	--	-----	8	10.00
Four-Year						
Private	13	16.25	5	6.25	18	22.50
Public	3	3.75	--	-----	3	3.75
University						
Private	9	11.25	--	-----	9	11.25
Public	27	33.75	15	18.75	42	52.50
TOTALS	60	75.00	20	25.00	80	100.00

Table III.3. Distribution of Composition Classes in Sample Across Freshman and Nonfreshman Courses by Institutional Type.

III. 3. RESULTS OF THE FACTOR ANALYSES

By using the two-part, 99-item questionnaire, this large group of students evaluated their composition courses and teachers. In evaluating their courses and teachers, they indicated the extent of their agreement or disagreement with the 99 statements. Agreement and disagreement were registered along a five-point Likert-type scale ranging from "strongly agree" to "strongly disagree." The scale's numerical values and their semantic equivalents are as follows: (1) strongly agree, (2) agree, (3) undecided, (4) disagree, and (5) strongly disagree.

If the student believed a particular statement or item was not relevant to his/her particular teacher or course, the student marked a sixth possible response--"not applicable to this course or this teacher."

The student responses to the 20 statements constituting the general part of the questionnaire were factor analyzed separately from the student responses to the 79 statements designed for evaluating the more specific aspects of composition teaching and courses. The factor analyses performed were principal component factor analyses with an orthogonal rotation as determined by Kaiser's varimax criterion (Kaiser, 1958; Morrison, 1967). All factor analyses were performed through use of the Statistical Analysis System (SAS) (Barr, Goodnight, Sall, & Helwig, 1976).

In deciding which items or statements loaded on which factors, we first invoked the 60-40/50-30 criterion (McCroskey & Young, 1979). According to this criterion, an item was judged to load satisfactorily on a factor if it met one of two conditions: (1) its numerical loading on a particular factor was 0.60 or greater and its numerical loadings on all other factors were no higher than 0.40, or (2) its numerical loading on one factor was at least 0.50 and its numerical loadings on all other factors were no greater than 0.30. If an item or statement met the 60-40/50-30 criterion, it was retained. If an item failed to meet the 60-40/50-30 criterion, we tried to determine whether it was tapping some important dimension of composition courses or composition teaching that was not tapped by other items. If an item seemed to be focusing on such a dimension, it was retained, even though it failed to meet the 60-40/50-30 criterion. In most instances, such items had factor loadings above 0.40; and in all but one case, all such items loaded primarily, although perhaps not strongly, on but one factor. Using this "liberal" policy for retaining statements, we hoped to remain true to the sources of the statements included on the questionnaire--the composition students, the composition teachers, and the writing program directors who had supplied either the statements themselves or the bases for the statements.

For both parts of the questionnaire, a number of factor solutions were generated and examined. For the 20-item general part of the instrument, we examined solutions having from two to nine factors. The two-factor solution was judged the most interpretable. This two-factor solution retained 18 of the original 20 items or statements and accounted for 96.1% of the variance in responses. For the 79-item second part, we examined solutions having from two to 19 factors. In this case, a nine-factor solution was determined to be the most interpretable. This nine-factor solution retained 62 of the original 79 items or statements and accounted for 91.2% of the variance in responses.

III. 3. 1. A Two-Factor Solution for Part I of the Questionnaire

The 18 items or statements retained in the two-factor solution for the general part of the questionnaire are listed in APPENDIX A. In APPENDIX A they are numbered from 1 through 18. A breakdown of the two-factor solution for the general part of the questionnaire appears in Table III.4. This table summarizes the "content" of the 18 statements retained by the two-factor solution, names the two factors, and lists the factor loadings for the 18 items. The factor loading of a particular item may be interpreted as the correlation of that item with what is being measured by the factor.

FACTORS/ITEMS	FACTOR LOADINGS
Factor 1: Teacher & Teaching	
1. good teacher	0.80
2. most helpful teacher	0.83
3. best teacher	0.75
4. teacher intellectually stimulating	0.70
5. instructional methods used	0.56
6. makes students think	0.53
7. treats students fairly	0.65
8. evaluates student's work fairly	0.67
*9. instructional methods help students become better writers	0.43
Factor 2: Course & Course Value	
*9. instructional methods help students become better writers	0.43
10. college-level writing course	0.51
11. compared to other courses requiring writing	0.49
12. assignments related to goals	0.44
13. instructional methods appropriate for type of course	0.41
14. usefulness now of what is taught	0.71
15. value of what is learned	0.87
16. teaches important things	0.81
17. usefulness in future of what is taught	0.83
18. what course does for student	0.45

Table III.4. Two-Factor Solution for Part I of the Instrument for Evaluating Teacher Effectiveness (*Item 9 loads equally on both factors).

The two factors listed in Table III.4 are general factors. The statements which loaded on the first factor all have to do with instruction, with either methods of instruction or the way the teacher conducts the course and treats the students. The statements associated with the first factor generally have relatively high factor loadings, suggesting that they are making strong contributions to the factor. The statements which loaded on the second factor, which we have called "Course and Course Value," focus less on teacher methods and behaviors and more on the nature of the course itself. Statements 14 through 18 seem to be tapping a sub-dimension which might be called "utility" or "usefulness." Except for the factor loading for Statement 18, those for this group are all higher than those for Statements 9 through 13. The latter statements, including Statement 9 which

loacs about ecually on toth factors, seem to treat directly or indirectly either the relationship of instructional methods to type of course or the value of the course as a writing course.

As was previously noted, the 18 items which loaded on the two factors for Part I of the questionnaire collectively explain 96.1% of the variance within all responses to Part I. The statements loading on Factor 1 (Teacher & Teaching) explain 49.6% of the variance, and the statements loading on Factor 2 (Course & Course Value) account for another 46.5% of the variance.

III. 3. 2. Coefficient Alpha Reliabilities for Part I

The coefficient alpha reliability (Cronbach, 1951) for Factor 1 is 0.89, for Factor 2, 0.90, and for all of Part I, 0.93. The coefficient alpha reliability provides a measure of the internal consistency of each of the two factors individually and of the scale as a whole. Coefficient alpha has a possible range from 0.00 (low reliability) to 1.00 (high reliability), with high values indicating that the items are contributing to the measurement of the same factor(s).

III. 3. 3. A Nine-Factor Solution for Part II of the Questionnaire

For the 79-item second and more specific part of the questionnaire, we decided that a nine-factor solution was the most interpretable. That solution for Part II, together with the number of statements retained for each of the nine factors and the amount of variance explained by each, appears in Table III.5.

FACTOR	# OF ITEMS RETAINED	% OF VARIANCE EXPLAINED
1. Aspects of Writing Instruction	19	24.9
2. Negative Aspects of Writing Instruction	10	12.4
3. Treating Student Writing Problems During Class	5	12.5
4. Encouraging Class Participation	4	8.5
5. Encouragement	10	11.8
6. Teaching Revision	5	8.0
7. Overemphasizing Grammar & Usage	3	5.3
8. Assigning, Commenting on, and Returning Student Writing	4	5.5
9. Teaching Students How to Draw on Personal Experience in Writing	2	2.3
TOTALS	62	91.2

Table III.5. Nine-Factor Solution for Part II of the Questionnaire. (Note: Factors are listed in the order of their extraction.)

As Table III.5 shows, of the original 79 statements or items included on the questionnaire, 62 are retained by the nine-factor solution. Collectively, the nine factors account for 91.2% of the variance within all responses to Part II of the questionnaire. Factor 1--which we have titled "Aspects of Writing Instruction"--accounts for 24.9% of the variance. Factor 2--called in Table III.5 "Negative Aspects of Writing Instruction"--explains another 12.4% of the variance; and Factor 3--"Treating Student Writing Problems During Class--accounts for 12.5%. The remaining six factors named in Table III.5 explain an additional 41.4% of the variance. These factors and the items loading on them are explained in more detail in the following paragraphs.

Factor 1: Aspects of Writing Instruction. Of the 62 items or statements retained by the nine-factor solution for Part II, 19 loaded on Factor 1, which explains 24.9% of the variance within all responses to Part II. We grouped these items into six clusters, each focusing on a different "aspect" of writing instruction. These six clusters, which were not determined by factor analyzing the statements loading on Factor 1, and the items associated with each are summarized in Table III.6. The complete statements summarized in Table III.6 appear in APPENDIX A.

CLUSTERS & STATEMENTS	FACTOR LOADINGS
Cluster 1: Types of Writing	
19. types of writing	0.64
20. connection between readings and types	0.62
Cluster 2: Models of Writing	
21. models of good writing	0.53
22. essays of professional writers	0.46
23. examples of writing	0.54
Cluster 3: Writing Purposes	
24. persuade audiences	0.73
25. argue thesis or position	0.76
26. inform audiences	0.70
Cluster 4: Writing Assignments	
27. specific about lengths of papers	0.28
28. help students write better in other classes	0.43
29. relation to what is taught	0.61
30. fit together	0.51
Cluster 5: Prewriting	
31. define issue to write about	0.57
32. explore writing topics	0.55
33. consider audience during planning	0.68
34. plan papers	0.59
Cluster 6: Organization & Development	
35. state main ideas	0.60
36. support main ideas	0.59
37. use logical reasoning	0.61

Table III.6. Breakdown and Summary of Statements Loading on Factor 1 of Part II.

Table III.6 subdivides the 19 statements loading on Factor 1 into six groups or clusters. These six clusters of items were

not themselves determined through a factor analysis of the statements loading on Factor 1. Rather these clusters were determined by carefully examining the 19 items individually and grouping those which appeared to be tapping the same aspect of composition teaching and courses. Thus the six clusters are conceptual groups; they may or may not be mathematical groups. With the exception of Statements 27 and 28--which have relatively low factor loadings of 0.28 and 0.43--the statements loading on Factor 1 represent a strong mathematical grouping of data for which we have supplied a general concept, "Aspects of Writing Instruction." The two statements, 27 and 28, with low factor loadings do not relate as strongly to the mathematical structure of Factor 1 as do the remaining 17 statements. However, they were so frequently mentioned by students as important dimensions to be considered that we decided to include them. In addition, those two statements loaded better on Factor 1 than on any of the other eight factors in the nine-factor solution.

Factor 2: Negative Aspects of Writing Instruction. Of the statements retained by the nine-factor solution for Part II, 10 were determined to load satisfactorily on Factor 2. These items appeared to us to fall into four categories or clusters. These clusters of statements are summarized in Table III.1. Complete statements appear in APPENDIX A.

CLUSTERS & STATEMENTS	FACTOR LOADINGS
Cluster 1: Evaluation of Student Writing	
38. too much emphasis on "usage"	0.42
39. too much emphasis on organization	0.59
40. comments often embarrass	0.54
41. grading standards too high	0.63
42. too few compliments	0.50
Cluster 2: Assignments	
43. boring assignments	0.53
44. not clearly related to goals	0.56
Cluster 3: Misuse of Class Time	
45. spends time inappropriately	0.63
46. lectures too often	0.62
Cluster 4: Teacher's Interest in Teaching	
47. not interested in teaching writing	0.53

Table III-7. Breakdown and Summary of Statements Loading on Factor 2 of Part II.

The 10 statements loading on Factor 2 account for 12.4% of the variance within the responses to Part II of the questionnaire. Of these 10 statements, five are associated with some aspect of the teacher's evaluation of student writing. A second cluster is formed by statements 43 and 44, both of which focus on the assignments teachers give. Two items, statements 45 and 46, can be categorized under "Misuse of Class Time." Statement 47 seems to fall into a category by itself and has to do with the teacher's interest in teaching writing. The clusters listed in Table III-7 were not identified by factor analyzing the items loading on Factor 2. The clusters were again formed by grouping statements conceptually.

Factor 3: Treating Writing Problems During Class. Five of the 62 statements loaded on Factor 3, which explains 12.5% of the variance within all responses to Part II of the questionnaire. Summaries of these statements together with their factor loadings appear in Table III-8.

STATEMENTS	FACTOR LOADINGS
48. helping with writing problems	0.61
49. encouraging discussion of writing problems	0.71
50. helping explore writing topics	0.74
51. helping as writing occurs	0.82
52. encouraging students to help one another	0.58

Table III.8. Summary of Statements Loading on Factor 3 of Part II.

All of the statements loading on Factor 3 focus on things the teacher does in class to help students with their writing. Statements 48 through 51 deal specifically with things the teacher does with students, while Statement 52 has to do with getting students to help one another. Except for Statement 49, the statements all seem to focus on "helping students" during the process of writing a paper.

Factor 4: Encouraging Class Participation. Four of the items retained in the nine-factor solution loaded on Factor 4, which we see as a "class participation" factor. Collectively, the four statements account for 8.5% of the variance within the responses to Part II of the instrument. The summaries of and the factor loadings for the statements appear in Table III.9.

STATEMENTS	FACTOR LOADINGS
53. encourages questions	0.77
54. responds positively to questions	0.50
55. encourages disagreement with teacher's views	0.50
56. encourages students to join discussions	0.78

Table III.9. Summary of Statements Loading on Factor 4 of Part II.

Factor 5: Encouragement. Ten items have been retained for Factor 5. Collectively, the statements explain 11.8% of the variance within all responses to Part II. These statements and their factor loadings are presented in Table III.10.

STATEMENTS	FACTOR LOADINGS
57. makes comments that encourage students to improve	0.33
58. is easy to talk to	0.44
59. takes students' writing problems seriously	0.52
60. tries to understand students' feelings about writing	0.65
61. makes students feel good about writing	0.80
62. makes students less afraid to write	0.84
63. tries to increase students' confidence	0.86
64. encourages students to do best writing	0.57
65. helps students enjoy writing	0.70
66. gets students to write for enjoyment	0.63

Table III.10. Summary of Statements Loading on Factor 5 of Part II.

The statements loading on Factor 5 all appear to reflect teacher behaviors which encourage students to write, which promote confidence in students about their writing, and which help students feel good about what they do in a writing course. Statements 57 and 58, which have lower factor loadings than the other eight statements, were retained because so many of the composition students we interviewed in generating the original set of 207 items stressed their importance. Five of the items with the highest factor loadings--Statements 61, 62, 63, 65, and 66--all seem to cluster around making the student feel good about writing. As is true for the other factors, Factor 5 represents a mathematical grouping of statements; and as is true of the other factors, the name resulted from our interpretation of the mathematical group.

Factor 6: Teaching Revision. Five statements have been retained for Factor 6, which explains 8.0% of the variance in Part II of the questionnaire. The factor loadings and summaries of these statements appear in Table III.11.

STATEMENTS	FACTOR LOADINGS
67. encouraging students to revise	0.68
68. teaching evaluation to improve revision	0.73
69. teaching how to revise organization	0.71
70. teaching to revise sentences	0.62
71. teaching to use teacher's comments in revision	0.55

Table III.11. Summary of Statements Loading on Factor 6 of Part II.

Two of the five statements focusing on revision--Statements 69 and 70--have to do with specific ends of revision, improving essay organization and sentences. Two statements--68 and 71--focus on teaching students how to use evaluations of the written product to revise that product. The remaining statement, 67, focuses on the encouragement teachers give to students to revise.

Factor 7: Overemphasizing Grammar & Usage. Three of the items focusing on negative aspects of writing instruction were grouped mathematically as a single factor, one which we have called "overemphasizing grammar and usage." As a single factor, these statements account for 5.3% of the variance. Summaries of these statements, together with their factor loadings, are presented in Table III.12.

STATEMENTS	FACTOR LOADINGS
72. puts too much emphasis on correct spelling	0.75
73. puts too much emphasis on correct grammar	0.79
74. puts too much emphasis on correct punctuation and capitalization	0.82

Table III.12. Summary of Statements Loading on Factor 7 of Part II.

Factor 8: Assigning, Commenting on, and Returning Student Papers. The four items retained for Factor 8 account for 5.5% of the variance within Part II. Summaries of the statements and their factor loadings appear in Table III.13.

STATEMENTS	FACTOR LOADINGS
75. establishes reasonable deadlines for papers	0.40
76. comments on papers are easy to read	0.55
77. comments on papers are easy to understand	0.56
78. getting evaluations of papers to students	0.29

Table III.13. Summary of Statements of Loading on Factor 8 of Part II.

These four statements all have to do with the teacher's handling of writing assignments. Two of the statements--75 and 78--have to do with time. Statement 75 was retained because of the importance the students we interviewed attached to it. Statement 75 focuses on the reasonableness of the deadlines the teacher establishes for assignments, and Statement 78 focuses on the amount of time between the submission and the return of the writing assignment. Statements 76 and 77 both deal with the ease with which students can assimilate the comments the teacher makes on completed assignments.

Factor 9: Teaching Students How to Draw on Personal Experience in Writing. The remaining statements make up Factor 9, which explains 2.3% of the variance. Both statements focus on the extent to which teachers encourage students to draw on their own feelings and experiences in writing, an aspect of teaching composition which is extremely important in some classes. Both items have relatively low factor loadings and were retained because they were the only statements which loaded above 0.18 on this factor. These statements and their factor loadings are presented in Table III.14.

STATEMENTS	FACTOR LOADINGS
79. encourages stucent to write about feelings	0.36
80. encourages student to write about personal experience	0.42

Table III.14. Summary of Statements Loacing on Factor 9 of Part II.

The nine-factor solution for Part II of the questionnaire reflects fairly well the conceptual structure of the original 207 items collected or gererated during Stage 1 of the questionnaire's development. The various dimensions of composition courses and instruction that writing teachers, students, and program directors said were important in evaluation of composition courses and teaching at the college level are all retained by the nine-factor solution. The principal difference between the 79 items retained for Part II and the original 207 items lies in their respective levels of specificity. For example, items among the original 207 focusing on discourse modes--e.g., description, classification, etc.--did load strongly on any factor in any of the factor solutions we examined.

III. 3. 4. Coefficient Alpha Reliabilities for Part II

Coefficient alpha reliabilities (Cronbach, 1951) were calculated for each of the factors in the nine-factor solution. Coefficient alpha reliabilities provide measures of the internal consistency of the nine factors individually. Such reliabilities can range from 0.00 (low reliability) to 1.00 (high reliability). High values indicate that the items are contributing to the measurement of the same factor(s). The coefficient alpha reliabilities for the nine factors of Part II of the long form are as follows: Factor 1--0.93; Factor 2--0.89; Factor 3--0.91; Factor 4--0.81; Factor 5--0.91; Factor 6--0.88; Factor 7--0.93; Factor 8--0.70; and Factor 9--0.78.

III. 4. CANONICAL CORRELATIONS BETWEEN THE TWO PARTS OF THE TEACHER AND COURSE EFFECTIVENESS INSTRUMENT

In this section, we report the results of a canonical correlation analysis used to determine the relationship between Part I and Part II of the long form of the questionnaire.

III. 4. 1. INTRODUCTION

It is frequently the case that instructors are unable to pinpoint or specifically explain the basic aspects of their instruction that may have been related to the ratings they receive from students in a particular course. For example, when below average or poor evaluations are received, the instructor is often at a loss in isolating specific steps that might be taken to correct the situation. The traditional instruments available for course and instructor evaluation provide insufficient guidance because they tend to supply only general or global information.

As we have stated, one of the major objectives guiding the development of the long form of the present instrument was to attempt to provide instructors with detailed information that could be used diagnostically to improve teaching. Certainly the rationale underlying the extraction of factors in Part II of the long form was to accomplish this objective. However, it is not enough simply to generate both specific items and general items, to divide them across two parts of a questionnaire, and then to assume the existence of strong relations between the two parts. Thus, a primary goal in using canonical correlation to analyze the long form of the instrument was to validate the hypothesized link between its two parts. Canonical correlation analysis was chosen for this investigation because it is a method specifically designed for determining correlations among two distinct sets of variables (Harris, 1975; Tatsuoka, 1971).

The reader will recall that the 18 items in Part I of the long form provided overall, global evaluations of the composition instructor and course, while the 62 items in Part II covered very detailed aspects of both the course and the instruction. Using canonical correlation analysis, we attempted to identify particular subsets or factors within the detailed items in Part II that have the strongest linear relationships with subsets of the general items or statements in Part I. The goal was thus to identify specific aspects of instruction and courses which correlate most highly with overall evaluations of courses and teachers. In the discussion that follows, however, the reader should bear in mind the axiom that correlation is insufficient to demonstrate causation.

III. 4. 2. Results of the Canonical Correlation Analysis

The results of the canonical correlation analysis are summarized in Table III.15 below.

Canonical Variate Pair	Canonical Corr.	Squares of Canon. Corr.	Asymptotic Chi-Square	DF	Prob > Chi-Sq.
1	0.90	0.81	3152.07	1116	.0001
2	0.59	0.34	1825.10	1037	.0001
3	0.50	0.25	1487.20	960	.0001
4	0.44	0.20	1256.39	885	.0001
5	0.42	0.17	1082.29	812	.0001
6	0.40	0.16	930.55	741	.0001

Table III.15. Statistically Significant Canonical Correlations Between Part I and Part II of the Long Form.

Using a very conservative level of statistical significance ($p < .0001$) to help balance the large sample size, six significant canonical correlations were found in the data. The canonical variate pairs (subsets of items from Part I that correlate with subsets of items from Part II) are listed in Table III.15. These canonical correlations range from 0.40 to 0.90. The column showing the "Squares of the Canonical Correlations" lists the proportion of variance in one canonical variate that is accounted for by the other canonical variate in the canonical variate pair.

The procedure for interpreting the results of canonical correlation analyses is similar to that used for factor analysis. Whereas factor loadings indicate the correlations of individual variables with a given factor, canonical loadings indicate the correlations of individual variables with one of the canonical variates in the pair. The basic procedure is to identify the variables with the highest canonical loadings and then to examine them conceptually to determine what they all seem to measure or what attributes they share. Based on the dimension they seem to be tapping, a descriptive label is assigned. This procedure is done separately for each canonical variate in each pair and across all pairs.

Table III.16 summarizes the results of the interpretation and classification of the canonical variates.

Canon. Var. Pair	Depend. Vars. (Items) w/ High Loadings	Depend. Canon. Var. Names	Canon. Corr.	Indepen. Vars. (Items) w/ High Loadings	Indepen. Canon. Var. Names
1	1 (.82) 2 (.78) 3 (.67) 4 (.71) 5 (.74) 7 (.73) 8 (.71) 9 (.72) 18(.75)	Teacher Effective-ness	0.90	19(.71) 57(.72) 60(.71) 61(.78) 62(.71) 63(.77) 64(.71) 65(.75)	Empathy & Encourage-ment
2	10(.32) 14(.52) 15(.58) 16(.50) 17(.53) 18(.30)	Course Utility	0.59	28(-.54) 53(.25) 55(.25) 58(.26)	Classroom Communica-tion
3	8 (.49)	Evalua-tion/ Fairness	0.50	39(-.31) 40(-.37) 41(-.54)	Criticism of Stu-dents' Writing
4	6 (.45) 9 (.32)	Student Develop-ment	0.44	22(.27) 61(-.29) 66(-.25)	Enjoy-ment
5	8 (.28) 16(.35)	Student Satisfac-tion	0.42	25(-.29) 27(-.31)	Speci-ficity
6	12(.31) 13(.33)	Assignments & Instr. Methods	0.40	55(.23) 80(.27)	Writing Confi-dence

Table III-16. Interpretations of Corresponding Canonical Variates Between Part I and Part II of the Long Form.

Each of the six pairs of canonical variates is listed, along with the subsets of items or statements with high canonical loadings, the magnitudes of the loadings indicated inside the parentheses. The items in Part I of the long form were designated as the dependent variables while those of Part II served as the independent variables. These designations were

arbitrary: either set could have been treated as dependent variables or incepencent variables, so long as they were treated consistently as one or the other. The label assigned to each element of the pairs of canonical variates denotes the dimension measured by each of the item subsets.

For example, within the first canonical pair, the nine statements or items from Part I provided a measure of teacher effectiveness. The eight items from Part II that are most highly correlated with teacher effectiveness comprised a measure of the empathy of the instructor with the students and the amount of encouragement the teacher provided. Looking at the second canonical variate pair, we see six items reflecting the students' ratings of the course's utility. These items are most highly correlated with the openness of communication in the classroom, as well as to the value of course content for other courses. The interested reader will wish to refer to APPENDIX A to inspect the individual statements listed in Table III.16 to determine the reasonableness of the labels assigned.

One of the major difficulties in the interpretation of canonical correlation results is that the correlations actually represent the degree of linear relationship between combinations of the variables rather than between the individual variables themselves. However, in the present analysis, most of the canonical variates were relatively easy to label because, conceptually, the items appear related in obvious ways. In other cases, compromises had to made and meanings stretched slightly to arrive at a descriptive label for a canonical variate. In general, the canonical correlation results provide satisfactory evidence for the validity of the presupposed link between Part I and Part II of the long form of the instrument. Hence, instructors should be able to use the detailed feedback provided by Part II of the questionnaire diagnostically, being assured that the specific items have a strong relationship with the more general evaluative items or statements of Part I.

III. 5. THE SHORT FORM OF THE COURSE AND TEACHER EFFECTIVENESS INSTRUMENT

The long form of the questionnaire provided the basis for the development of a 21-item short form. We developed this short form to provide teachers with a more easily scored, reliable, and valid instrument that could be used in two situations: (1) when time precludes administering the 80-item long form and (2) when the detailed information elicited by the long form is not required.

III. 5. 1. Development of the Short Form

The 21-item short form was developed after we had determined

the factor structures of Part I and Part II of the long form. With the factor solutions of Part I and Part II of the long form at hand, we selected those items or statements which seemed to reflect best the factor solutions of the long form. The statements selected for inclusion on the short form were typically those with the highest factor loadings. Given the sources of the statements, their content validity was assumed.

For the short form, we selected seven items from Part I of the long form, four loading on Factor 1 and three loading on Factor 2. We selected the statements from Part I on the basis of two considerations: (1) whether they had relatively high factor loadings and (2) whether they seemed to be focusing on some aspect of instruction and courses which students, teachers, and writing program directors had said were especially important. Then we selected 14 items from Part II of the long form. Six of these 14 statements loaded on Factor 1 of Part II, one representing each of the six clusters presented in Table III.6. Two of the 14 statements loaded on Factor 2 of Part II, one representing Cluster 1 and one representing Cluster 3 as depicted in Table III.7. Of the remaining six statements selected for inclusion on the short form of the questionnaire, one was from those loading on each of the remaining factors except Factor 9. Factors 3 through 8 of the Part II are summarized in Table III.8 through III.13 in the previous section. In selecting which statements from Part II to include on the short form, we used the same two criteria we had used in selecting items from Part I. A statement loading on Factor 9 of Part II was not included on the short form because both of the possible selections had low factor loadings (0.36 and 0.42, respectively) and because Factor 9 of Part II accounted for only 2.3% of the variance.

III. 5. 2. A Two-Factor Solution for the Short Form

The 21-item short form was submitted to factor analysis. The results of this analysis and the relationship between the short and long forms of the questionnaire are presented in Table III.17. The short form of the questionnaire appears in APPENDIX B.

SHORT FORM			LONG FORM		
Item #	Factor	Factor Loadings	Item #	Part & Factor	Factor Loadings
1	1	0.72	1	I-1	0.80
2	1	0.74	2	I-1	0.83
3	1	0.63	4	I-1	0.70
4	1	0.66	8	I-1	0.67
5	1	0.64	19	II-1.1	0.64
6	1	0.54	25	II-1.3	0.76
7	1	0.13	73	II-7	0.79
8	1	0.55	29	II-1.4	0.61
9	1	0.51	68	II-6	0.73
10	1	0.52	33	II-1.5	0.68
11	1	0.48	51	II-3	0.82
12	1	0.44	45	II-2.3	0.63
13	1	0.43	41	II-2.1	0.63
14	1	0.58	77	II-8	0.56
15	1	0.61	36	II-1.6	0.59
16	1	0.53	23	II-1.2	0.54
17	1	0.54	56	II-4	0.76
18	1	0.73	63	II-5	0.86
19	2	0.71	14	I-2	0.71
20	2	0.86	15	I-2	0.87
21	2	0.79	17	I-2	0.83

Table III.17. Relationship Between the Short and Long Forms of the Questionnaire.

As Table III.17 shows, we decided on a two-factor solution for the short form of the questionnaire. The two-factor solution was the only interpretable solution of the several examined. Both factors are general factors. Factor 1 of the short form we call "Instruction." Of the 21 items included on the short form, 18 load on this factor, although the factor loading of item 7 is fairly low. Factor 1--"Instruction"--accounts for 66.7% of the variance. The second factor, which we call simply "Content," accounts for an additional 29.1% of the variance. The three statements loading on Factor 2 all focus on the value of *what* the course teaches, that is, the value of its *content* or *curriculum*. In contrast, the 18 items loading on Factor 1--"Instruction"--all focus on *how* and *how well* the teacher teaches that content. From a conceptual point of view, the two-factor solution for the short form may be thus said to focus on *instruction* and *curriculum*. Collectively, these two factors explain 95.8% of the variance of the short form.

III. 5. 3. Coefficient Alpha Reliabilities for the Short Form of the Instrument

Coefficient alpha reliabilities (Cronbach, 1951) were computed for the short form to provide a measure of the internal consistency of the two factors individually and of the short form as a whole. The alpha coefficient reliability for the short form is 0.86. For Factor 1--"Instruction"--it is 0.84; and for Factor 2--"Content"--it is also 0.84.

CHAPTER IV

THE SCORING, USES, AND LIMITATIONS OF THE INSTRUMENT

IV. 1. INTRODUCTION

The present brief chapter focuses on issues which should be of interest to persons and institutions that may be considering using either the long or the short form of the questionnaire in their composition programs. The chapter has three major parts: (1) scoring the long and short forms of the instrument; (2) recommended uses of the instruments; and (3) the limitations of the instruments.

IV. 2. SCORING THE LONG AND SHORT FORMS

At the outset of this report, we said that one of our purposes was to develop an instrument for measuring the effectiveness of composition courses and teaching at the college level. We stated that such an instrument should be sensitive to the specific demands of the discipline. The two parts of the long form of the instrument yield evaluations of composition courses and teaching along 11 different conceptual and mathematical dimensions or factors, while the short form yields evaluations along two dimensions or factors. Regardless of which form is used, scoring should reflect the various dimensions of composition courses and teaching measured; only through such scoring can the instruments yield the maximally specific information teachers may need to improve their teaching.

In scoring the long form of the questionnaire, we recommend the following procedures. It will be recalled that a response to a given item could be either "strongly agree," "agree," "undecided," "disagree," or "strongly disagree" with numerical values ranging from 1 to 5. Most of the items included on the long form are stated positively. For these items, a response of 1 ("strongly agree") would be a favorable one. However, 13 of the items or statements (Statements 38 through 47 and 72 through 74) are worded negatively. A response of 1 ("strongly agree") to these 13 items would be an unfavorable one, and a response of 5 ("strongly disagree") would be a favorable one. Thus in scoring the responses to the 13 negative items, the 1-to-5 scale needs to be reversed, an especially simple procedure if the student responses are computer-scored.

Once the scale is reversed for the 13 negative items, the number of student responses to each of the items associated with

ary one factor should be summed across the statements for that factor. These sums should then be divided by (1) the number of statements making up that factor and (2) the number of students. These divisions will yield an average score ranging from 1 to 5, where a score close to 1.0 is highly favorable and a score close to 5.0 is highly unfavorable. The same procedure should be repeated for each factor. It is not recommended that separate scores be computed for the "clusters" of items associated with Factor 1 and Factor 2 of Part II because these clusters do not represent mathematical clusters, only conceptual ones.

Scoring the student responses to the long form of the questionnaire will result in profiles of performance in composition teaching and courses. Each profile generated through the scoring of the long form will consist of at least 11 subscores, one for each of the two factors of Part I and one for each of the nine factors of Part II.

Similar procedures can be used to compute three additional scores, scores which do not reflect the factor structures of the two parts of the long form. A composite score can be computed for Part I and for Part II, and an overall score can be computed for the entire long form. None of the computations is difficult, but each is likely to be time-consuming unless performed by a computer. It should be pointed out that the long form was developed and the factor structures identified in order to provide teachers with specific information about their teaching and their courses. Composite scores for Part I and for Part II and an overall score for the entire long form will yield only general information. We do not recommend using or even computing a total overall score for the long form or for Part II of the long form because doing so contradicts the purpose for which Part II was developed.

The short form of the questionnaire can be similarly scored. In this case, the 5-point scale needs to be reversed for three negatively worded items--7, 12, and 13. Separate scores should be computed for each of the two factors of the short form. The score for the first factor would represent a general estimate of the instruction, and the score for Factor 2 would provide a general estimate of the course content (curriculum) itself. A total score for the short form would yield an overall, global evaluation of the teaching and the course. The scores for the short form will provide the teacher with less specific information than will the subscores for the 11 factors of the long form.

As a general rule, once the 5-point scale is reversed for the negative items, mean scores smaller than 3 will indicate positive evaluations. The closer the mean score is to 1, the more positive the evaluation.

IV. 3. RECOMMENDED USES OF THE INSTRUMENT

We developed the instrument in its two forms to provide information to teachers about their teaching and their courses. Either form of the instrument can be used in either formative or summative evaluations. Because of the large number of items contained on it, we recommend that the long form be used only in summative evaluations. The short form of the instrument and Part I of the long form, on the other hand, can be easily used for either type of evaluation. Which form is used by a given teacher should depend on (1) how much information the teacher wants about his/her course and teaching and (2) how much of the students' time the teacher wishes to devote to evaluation.

Although the two forms of the instrument were designed specifically to provide information to teachers, we suspect that they will be used for other purposes as well. We anticipate that administrators will use the resultant scores to make decisions regarding raises, promotion, and tenure. For this purpose the short form is better suited than the long form. Given the highly significant canonical correlations between the two parts of the long form, Part I of the long form is a suitable alternative for the short form. Whether the short form or Part I of the long form is used for this purpose, only "factor" scores or total scores should be submitted to administrators. Never should administrators be given scores for individual statements: such statements are meaningful only in the context of the factor on which they load. We see no situation where an administrator should have access to the responses to Part II of the long form for the purpose of making decisions regarding raises, promotion, or tenure. The information yielded by Part II of the long form is diagnostic and specific in nature, and should be used as a basis for self-evaluation of teaching, not as a basis for an administrative decision about the worth of a particular teacher.

We also anticipate that some persons will want to make available the scores generated by the two forms of the instrument as an aid to students in choosing teachers. Although neither form of the instrument was designed specifically for that purpose, the short form and Part I of the long form are much better suited to that purpose than Part II of the long form or the long form as a whole. In fact, we believe that the specific, diagnostic information provided by Part II of the long form should not be made available to students. As in the case of administrators, scores for individual statements should not be made available to students. Only "factor" scores and/or total scores for the short form or Part I of the long form should be so "published." Of course, the statements whose scores constitute the "factor" scores should be made available, but not the responses to or scores for individual statements.

IV. 4. CAUTIONS REGARDING THE USE OF THE INSTRUMENTS

The two forms of the instrument we have described in the

present document have been very carefully developed over a period of nearly 20 months. Because of the way they were developed and because they were developed specifically for use in college composition courses, they are probably the best such instruments available for evaluating college composition teaching and courses. They meet all the criteria we set for such instruments at the beginning of this report: they have content validity; they are reliable; and they can be easily scored. Yet certain cautions regarding their use are in order.

Any user of the forms of the instrument should be aware of the hazards of any self-report instrument. No matter how much time and effort go into the development of a self-report instrument, such instruments can be no better than the responses they elicit. Such instruments can not control for responses from students who are "out to get" an instructor for whatever reason; they can not control for responses from students who mindlessly mark possible responses; and they cannot control for negative or positive responses which are biased by influences beyond the control of the teacher or which lie beyond the scope of the course.

Interpreting the various scores should also be done cautiously. At the present we do not have sufficient data or file to report norms for different types of writing courses. We need a data base for a large variety of college writing courses, courses ranging from developmental or basic writing courses to advanced undergraduate courses. When such a data base is constructed, we will be able to supply users of the two forms the kind of normative information that will allow them to compare their teaching and their courses with those of others in similar settings. For the most part, the two forms were developed from responses gathered in freshman writing courses (see Chapter II for a description of the sample population).

A third caution has also to do with the interpretation and use of the scores. Neither of the two forms should be used as the <u>only</u> source of evaluation data. The instruments represent but one kind of evaluation, albeit the most commonly used one. The type of evaluation which the present instruments represent should be supplemented by other types, such as evaluation by fellow faculty, evaluations of course material and syllabi, self-evaluation, delayed evaluations by students, and evaluations of student performance in subsequent courses. In addition, dimensions of teaching and courses beyond those we have identified also need to be evaluated--the amount and kind of preparation the teacher does, the adequacy of the curriculum for the students served, the relationship of a particular class to the writing program, and so forth. We would hope that while the two forms of the instrument we have developed can distinguish between effective and ineffective teachers and courses, other means of evaluation will be employed before important decisions are made regarding either teachers or courses.

Finally, there is no way we could build into the forms of

the instrument checks against the misuse of information they supply. We anticipate that the most flagrant misuses will occur when the information supplied by the instruments is used to make administrative decisions. We anticipate that chairpersons and committees alike will, for example, use the mean response for <u>one or two items</u> to make decisions about particular teachers. Such an item is Statement 11 on Part I of the long form: "Compared with other courses which require students to write . . . , this course is a very good one." Such an item is meaningful only in the context of the other items loading on the same factor; no single item can be interpreted out of its factorial context without violating the underlying structure of the instrument itself. Once such a violation occurs, the response can no longer be considered either reliable or valid.

The cautions we have issued, of course, reflect as much the misuse of such questionnaires as they do the inherent weaknesses of the instruments we have developed. We have tried to develop instruments worthy of our fellow teachers and worthy of their students. If used properly and interpreted correctly, the information supplied by the instruments should help good teachers of writing to become better ones.

REFERENCES

Barr, A.J., J.H. Goodnight, J.P. Sall, and J.T. Helwig. *A user's guide to SAS 76*. Raleigh, NC: SAS Institute, 1976.

Biesel, C.E., et al. (eds). *The college blue book: narrative description*, 17th ed. New York: Macmillan, 1979.

Cass, J., and Birnbaum, M. *Comparative guide to American colleges*, 5th ed. New York: Harper & Row, 1979.

Centra, J.A. The how and why of evaluating teaching. In J.A. Centra (ed.), *Renewing and evaluating teaching*. San Francisco: Jossey-Bass, 1977.

Centra, J.A. *How universities evaluate faculty performance: a survey of department heads*. GRE Board Research Report No. 75-5bR. Princeton, NJ: Educational Testing Service, 1977.

Centra, J.A. The student as godfather? the impact of student ratings on academia. *Educational Researcher*, 1973, *2*, 4-8.

Costin, F., W.T. Greenough, and R.J. Menges. Student ratings of college teaching: reliability, validity, and usefulness. *Review of Educational Research*, 1971, *XX*, 511-535.

Cromack, T.R., and Egelston, R.L. Students as the source of information concerning teacher behaviors: new concepts in the appraisal of instruction. *Journal of Classroom Interaction*, 1979, *15*, 8-15.

Cronbach, L.J. Coefficent alpha and the internal structure of tests. *Psychometrika*, 1951, *16*, 297-334.

Dearman, N.B. *The condition of education: statistical report*. Washington, D.C.: National Center for Education Statistics, 1980.

Eble, K. *The recognition and evaluation of teaching*. Washington, D.C.: American Association of University Professors and the Association of American Colleges, 1970.

Frey, P.W. The ongoing debate: student evaluation of teaching. *Change*, Feb., 1974, 47-48, 64.

Grant, W.V., and L.J. Eiden. *Digest of education statistics*. Washington, D.C.: National Center for Education Statistics, 1980.

Harris, R.J. A primer of multivariate statistics. New York: Academic Press, 1975.

Kaiser, H.F. The varimax criterion for analytic rotation in factor analysis. Psychometrika, 1958, 23, 187-200.

Kaiser, H.F. Little Jiffy, Mark IV. Educational and Psychological Measurement, 1974, 34, 111-117.

Kim, J-O., and C.W. Mueller. Factor analysis: statistical methods and practical issues. Quantitative Applications in the Social Sciences, 14. Beverly Hills and London: Sage Publications, 1978a.

Kim, J-O., and C.W. Mueller. Introduction to factor analysis: what it is and how to do it. Quantitative Applications in the Social Sciences, 13. Beverly Hills and London: Sage Publications, 1978b.

Kulik, J.A. Feed to the faculty: student reactions to instruction. Center for Research on Learning and Teaching, University of Michigan, 1976.

Kulik, J.A., and W.J. McKeachie. The evaluation of teachers in higher education. In F.N. Kerlinger (ed.), Review of educational research, vol. 3. Itasca, IL: F.E. Peacock, 1975.

Larson, R.L., and the CCCC Committee on Teaching and Its Evaluation in Composition. Evaluating instruction in composition: approaches and instruments. Unpublished ms.

McCroskey, J.C., and T.J. Young. The use and abuse of factor analysis in analysis in communication research. Human Communication Research, 1979, 5, 375-382.

Marsh, H.W., and J.L. Overall. Validity of students' evaluation of teaching effectiveness: cognitive and affective criteria. Journal of Educational Psychology, 1980, 72, 468-475.

Morrison, D.F. Multivariate statistical methods. New York: McGraw-Hill, 1967.

Newton, R.R. Teacher evaluation: focus on outcomes. Peabody Journal of Education, 1980, 58, 45-54.

O'Hanlon, J., and L. Porterson. Making teacher evaluation work. Journal of Higher Education, 1980, 51, 664-672.

Powell, R. Faculty rating scale validity: the selling of a myth. College English, 1978, 39, 616-629.

Shock, N.W., E.L. Kelly, and H.H. Remmers. An empirical study of the validity of the Spearman-Brown formula as applied to

the Purdue Rating Scale. Journal of Educational Psychology, 1927, 18, 187-195.

Tatsuoka, M.M. Multivariate analysis. New York: John Wiley & Sons, 1971.

Thorne, G.L. Student ratings of instructors: from scores to administrative decisions. Journal of Higher Education, 1980, 51, 207-214.

Tuchman, B.W. The Tuchman Teacher Feedback Form (TTFF). Journal of Educational Measurement, 1976, XX, 233-237.

Tuchman, B.W., and W.S. Oliver. Effectiveness of feedback to teachers as a function of source. Journal of Educational Psychology, 1968, 59, 297-301.

Watts, S.F., and J. Hunter (eds.). Peterson's annual guide to undergraduate study, 1980 edition. Princeton, NJ: Peterson's Guides, 1979.

Whitely, S.E., and K.O. Doyle. Validity and generalizability of student ratings from between-class and within-class data. Journal of Educational Psychology, 1979, 71, 117-124.

Witte, S.P., C.A. Daly, and L. Faigley. Assessing the teaching of college writing. Paper presented during the Annual Convention of the National Council of Teachers of English. Cincinnati, Nov., 1980.

Witte, S.P., P.F. Meyer, T.P. Miller, and L. Faigley. A national survey of college and university writing program directors. FIPSE Grant No. G008005896 Technical Report No. 2. Austin: The University of Texas at Austin, Writing Program Assessment Office, 1981.

APPENDIX A

INSTRUMENT FOR EVALUATING COURSE AND TEACHER EFFECTIVENESS: LONG FORM

Copyright c 1981 by Stephen P. Witte

Part I

1. My teacher in this course is a very good one.

2. My teacher in this course is one of the most helpful teachers I have ever had.

3. My teacher in this course is the best teacher I have ever had.

4. The instructor is intellectually stimulating.

5. The instructional methods (such as lectures, student conferences, class discussions, homework assignments) which are used in this course are very effective.

6. The instructor is good about making students think for themselves.

7. As a student in this course, I am treated fairly.

8. My work as a student in this course is evaluated fairly.

9. The instructional methods which are used in this courses are very good in helping students to become better writers.

10. This course is an excellent college-level writing course.

11. Compared with other college courses which require students to write (courses such as history courses, political science courses, and geography courses), this course is a very good one.

12. The assignments (such as exercises, essays and

papers, and readings) which are made in this course are closely related to the goals of the course.

13. The instructional methods which are used in this course are the right ones for the type of course it is.

14. What this course teaches is very useful to the student right now.

15. What I am learning in this course is valuable.

16. This course teaches important things that every college graduate should know.

17. What this course teaches will be useful to its students in the future.

18. As a student in this course, I feel good about what it has done for me personally.

Part II

19. The instructor does a good job of teaching students how to write different types of papers.

20. The instructor is good about making connections between assigned readings and the types of writing students do in this class.

21. The instructor is good about providing students with models of good writing.

22. The instructor is good about using class time to discuss essays that professional writers have written.

23. The instructor does a good job of using examples of writing in teaching students how to write.

24. The instructor is good about teaching students how to persuade audiences through writing.

25. The instructor is good about teaching students how to argue a thesis or a position.

26. The instructor is good about teaching students how to inform audiences through writing.

27. The instructor is specific enough about the length of the papers students are assigned to write.

28. The writing assignments I did in this course will probably help me write better in my other classes.

29. The instructor clearly connects what he or she teaches in class with the writing assignments he or she makes.

30. The instructor is good about making sure the writing assignments fit together.

31. The instructor does a good job of teaching me how to define an issue to write about.

32. The instructor does a good job of teaching me how to explore a writing topic.

33. The instructor is good about teaching me to consider the needs of my audience when I am thinking about a writing topic.

34. The instructor is good about teaching me how to plan my papers before I write them.

35. The instructor is good about teaching me how to state main ideas in my papers.

36. The instructor is good about teaching me how to support main ideas in my papers through examples and specific details.

37. The instructor is good about teaching me how to use logical reasoning in the papers I write.

38. In evaluating my writing, the instructor places too much emphasis on spelling, punctuation, grammar, and capitalization and not enough on my ideas.

39. In evaluating my writing, the instructor places too much emphasis on the way I have organized my paper.

40. The instructor's comments on my papers too often embarrass me about my writing.

41. In evaluating my writing, the instructor uses standards that are too high.

42. In commenting on my papers, the instructor compliments me very little when I write well.

43. The instructor's writing assignments are boring.

44. The instructor's writing assignments are not clearly related to the goals of the course.

45. The instructor spends too much time on some

things anc not enough time on others.

46. The teacher lectures tco ofter.

47. The irstructor seems less interested in teaching writing than he or she shculc be.

48. During class periocs, the instructor is gocd about helping me with my writing problems.

49. The instructor is good about encouraging me to discuss my writing problems in class.

50. The irstructor is good about using class time to help me explore topics for writing.

51. The irstructor is good about using class time to help me as I am writing my papers.

52. The instructor is good about encouraging me to use class time to help other student with the papers they're writing.

53. The instructor is gocd about encouraging me to ask cuestions ir class.

54. The instructor is gooc about responding positively tc questions I ask ir class.

55. The instructor is gocd about encouraging me to disagree with what he or she says.

56. The instructor is good about enccuraging me to join in class ciscussions.

57. In evaluating my writing, the instructor is gocd about making comments that encourage me tc beccme a better writer.

58. The instructor is easy tc talk to.

59. The instructor is good about taking my writing problems sericusly.

60. The irstructor is good about trying to understand the feelings I have about my writing.

61. The instructor does a good job making me feel gooc about my writing.

62. The instructor is good about making me feel less afraic tc write.

63. The instructor is good about trying to increase my confidence about writing.

64. The instructor does a good job encouraging me to do my best writing.

65. The instructor is good about making me enjoy writing.

66. The instructor does a good job of getting me to write for enjoyment rather than because I have to.

67. The instructor is good about encouraging me to rewrite or revise my papers.

68. The instructor is good about teaching me to evaluate my own papers so that I can better revise them.

69. The instructor is good about teaching me how to revise the organization of the papers I write.

70. The instructor is good about teaching me to revise my sentences so that they become more effective.

71. The instructor is good about teaching me how to use his or her comments about my papers as I revise.

72. The instructor puts too much emphasis on correct spelling in student writing.

73. The instructor puts too much emphasis on "correct" grammar in student writing.

74. The instructor puts too much emphasis on acceptable punctuation and capitalization in student writing.

75. The instructor establishes reasonable deadlines for the papers students write in his or her course.

76. The instructor's comments on my papers are easy to read.

77. The instructor's comments on my papers are easy to understand.

78. The instructor is good about getting his or her evaluations of my papers to me within a reasonable time.

79. The instructor is good about encouraging students to write about their own feelings.

80. The instructor is good about giving writing assignments that encourage students to draw on personal experience.

APPENDIX B

INSTRUMENT FOR EVALUATING COURSE AND TEACHER EFFECTIVENESS: SHORT FORM

Copyright © 1981 by Stephen P. Witte

1. My teacher in this course is a very good one.

2. My teacher in this course is one of the most helpful teachers I have ever had.

3. The instructor is intellectually stimulating.

4. My work as a student in this course is evaluated fairly.

5. The instructor does a good job of teaching students how to write different types of papers.

6. The instructor is good about teaching students how to argue a thesis or position.

7. The instructor puts too much emphasis on "correct" grammar in student writing.

8. The instructor clearly connects what he or she teaches in class with the writing assignments he or she makes.

9. The instructor is good about teaching me to evaluate my own papers so that I can better revise them.

10. The instructor is good about teaching me to consider the needs of my audience when I am thinking about a writing topic.

11. The instructor is good about using class time to help me as I am writing my papers.

12. The instructor spends too much time on some things and not enough time on others.

13. In evaluating my writing, the instructor uses standards that are too high.

14. The instructor's comments on my papers are easy to understand.

15. The instructor is good about teaching me how to support main ideas in my papers through examples and specific details.

16. The instructor does a good job of using examples of writing in teaching students how to write.

17. The instructor is good about encouraging me to join in class discussions.

18. The instructor is good about trying to increase my confidence about writing.

19. What this course teaches is very useful to the student right now.

20. What I am learning in this course is valuable.

21. What this course teaches will be useful to its students in the future.